

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: Yeda Research and Development Co. Ltd.
(B) STREET: Weizmann Institute of Science
(C) CITY: P.O.B. 95
(D) STATE: Rehovot
(E) COUNTRY: Israel
(F) POSTAL CODE (ZIP): 76100
(G) TELEPHONE: +972-8-9344093
(H) TELEFAX: +972-8-9470739

(A) NAME: David Wallach
(B) STREET: 24 Borochoy Street
(C) CITY: Rehovot
(E) COUNTRY: Israel
(F) POSTAL CODE (ZIP): 76406

(A) NAME: Nikolai Malinin
(B) STREET: Beit Clore, Weizmann Institute of Science
(C) CITY: Rehovot
(E) COUNTRY: Israel
(F) POSTAL CODE (ZIP): 76100

(A) NAME: Mark Boldin
(B) STREET: Beit Clore, Weizmann Institute of Science
(C) CITY: Rehovot
(E) COUNTRY: Israel
(F) POSTAL CODE (ZIP): 76100

(A) NAME: Andrei Kovalenko
(B) STREET: Beit Clore, Weizmann Institute of Science
(C) CITY: Rehovot
(E) COUNTRY: Israel
(F) POSTAL CODE (ZIP): 76100

(A) NAME: Igor Mett
(B) STREET: 60 Levin Epstein Street
(C) CITY: Rehovot
(E) COUNTRY: Israel
(F) POSTAL CODE (ZIP): 76462

(ii) TITLE OF INVENTION: Modulators of TNF Receptor Associated Factor (TRAF), their Preparation and Use

(iii) NUMBER OF SEQUENCES: 11

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(v) CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/IL97/00117

(vi) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: IL 117800
(B) FILING DATE: 02-APR-1996

(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: IL 119133
(B) FILING DATE: 26-AUG-1996

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1906 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CATTGGGTCA	CGCGGTGGCG	GCGCTCTAGA	ATAGTGGATC	CCCCGGGCTG	CAGGAATTCG	60
ATTCGAGGCC	ACGAAGGCCG	GCGGCGCGGC	GCANGCACCG	CCCCGGGGAN	AGGCNCCATG	120
AGCGGATCNC	NGAACNATGA	CAAAAGACAA	TTTCTGCTGG	AGCGACTGCT	GGATGCAGTG	180
AAACAGTGCC	AGATCCGCTT	TNGAGGGAGA	AAGGAGATTG	CCTCGGATTC	CGACAGCAGG	240
GTCACCTGTC	TGTGTGCCCA	GTTTGAAGCC	GTCCTGCAGC	ATGGCTTGAA	GAGGAGTCGA	300
GGATTGGCAC	TCACAGCGGC	AGCGATCAAG	CAGGCAGCGG	GCTTTGCCAG	CAAAACCGAA	360
ACAGAGCCCCG	TGTTCTGGTA	CTACGTGAAG	GAGGTCTCTA	ACAAGCACGA	GCTGCAGCGC	420
TTCTACTCCC	TGCGCCACAT	CGCCTCAGAC	GTGGGCCGGG	GTCGCGCCTG	GCTGCGCTGT	480
GCCCTCAACG	AACACTCCCT	GGAGCGCTAC	CTGCACATGC	TCCTGGCCGA	CCGCTGCAGG	540
CTGAGCACTT	TTTATGAAGA	CTGGTCTTTT	GTGATGGATG	AAGAAAGGTC	CAGTATGCTT	600
CCTACCATGG	CAGCAGGTCT	GAACTCCATA	CTCTTTGCGA	TTAACATCGA	CAACAAGGAT	660
TTGAACGGGC	AGAGTAAGTT	TGCTCCCACC	GTTTCAGACC	TCTTAAAGGA	GTCAACGCAG	720
AACGTGACCT	CCTTGCTGAA	GGAGTCCACG	CAAGGAGTGA	GCAGCCTGTT	CAGGGAGATC	780
ACAGCCTCCT	CTGCCGTCTC	CATCCTCATC	AAACCTGAAC	AGGAGACCGA	CCCTTGCCCTG	840
TCGTGTCCAG	GAATGTCAGT	GCTGATGCCA	AATGCAAAAA	GGAGCGGAAG	AAGAAAAAGA	900
AAGTGACCAA	CATAATCTCA	TTTGATGATG	AGGAAGATGA	GCAGAACTCT	GGGGACGTGT	960
TTAAAAAGAC	ACCTGGGGCA	GGGGAGAGCT	CAGAGGACAA	CTCCGACCGC	TCCTCTGTCA	1020
ATATCATGTC	CGCCTTTGAA	AGCCCCCTTCG	GGCCTAACTC	CAATGGAATC	AGAGCAGCAA	1080
CTCATGGAAA	ATTGATTCCC	TGTCTTTGAA	CGGGGAGTTT	GGGTACCAGA	AGCTTGATGT	1140
GAAAAGCATC	GATGATGAAG	ATGTGGATGA	AAACGAAGAT	GACGTGTATG	GAAACTCATC	1200

AGGAAGGAAG CACAGGGGCC ACTCGGAGTC GCCCGAGAAG CCACTGGAAG GGAACACCTG 1260
 CCTCTCCCAG ATGCACAGCT GGGCTCCGCT GAAGGTGCTG CACAATGACT CCGACATCCT 1320
 CTTCCCTGTC AGTGGCGTGG GCTCCTACAG CCCAGCAGAT GGGGGGCTCG GAAGCCTGGA 1380
 GAACGGGACA GGACCAGAGG ACCACGTTCT CCCGGATCCT GGACTTCGGT ACAGTGTGGA 1440
 AGCCAGCTCT CCAGGCCACG GAAGTCCTCT GAGCAGCCTG TTACTTCTGC CTCAGTGCCA 1500
 GAGTCCATGA CAATTAGTGA ACTGCGCCAG GCCACTGTGG CCATGATGAA CAGGAAGGAT 1560
 GAGCTGGAGG AGGAGAACAG ATCACTGCGA AACCTGCTCG ACGGTGAGAT GGAGCACTCA 1620
 GCCGCGCTCC GGCAAGAGGT GGACACCTTG AAAAGGAAGG TGGCTGAACA GGAGGAGCGG 1680
 CAGGGCATGA AGGTCCAGGC GCTGGCCAGC TATCTTTGCT ATTTTGTGAG GAGATTCTAA 1740
 CCCCACGTGA GAACCATGTG GTGGAGAAAT GGAGGGAGAG AGAAATCCAA CAGTTCCTGA 1800
 TAGTCTCATT TGAGCTCCTG GATCCAGTCT TTCCTGAAGC TGTGTTTCCT CTGGACTTTT 1860
 CATGTATGTG AGCCAATAAA TTGCTTTCAT TCCTTGAAAA AAAAAA 1906

(2) INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 604 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Xaa	Thr	Gly	Pro	Gly	Xaa	Gly	Xaa	Met	Ser	Gly	Ser	Xaa	Asn	Xaa	Asp
1				5					10					15	
Lys	Arg	Gln	Phe	Leu	Leu	Glu	Arg	Leu	Leu	Asp	Ala	Val	Lys	Gln	Cys
			20					25					30		
Gln	Ile	Arg	Phe	Xaa	Gly	Arg	Lys	Glu	Ile	Ala	Ser	Asp	Ser	Asp	Ser
		35					40					45			
Arg	Val	Thr	Cys	Leu	Cys	Ala	Gln	Phe	Glu	Ala	Val	Leu	Gln	His	Gly
	50					55					60				
Leu	Lys	Arg	Ser	Arg	Gly	Leu	Ala	Leu	Thr	Ala	Ala	Ala	Ile	Lys	Gln
65					70				75					80	
Ala	Ala	Gly	Phe	Ala	Ser	Lys	Thr	Glu	Thr	Glu	Pro	Val	Phe	Trp	Tyr
			85					90					95		
Tyr	Val	Lys	Glu	Val	Leu	Asn	Lys	His	Glu	Leu	Gln	Arg	Phe	Tyr	Ser
		100						105					110		
Leu	Arg	His	Ile	Ala	Ser	Asp	Val	Gly	Arg	Gly	Arg	Ala	Trp	Leu	Arg

Cys	Ala	Leu	Asn	Glu	His	Ser	Leu	Glu	Arg	Tyr	Leu	His	Met	Leu	Leu
130						135					140				
Ala	Asp	Arg	Cys	Arg	Leu	Ser	Thr	Phe	Tyr	Glu	Asp	Trp	Ser	Phe	Val
145					150					155					160
Met	Asp	Glu	Glu	Arg	Ser	Ser	Met	Leu	Pro	Thr	Met	Ala	Ala	Gly	Leu
				165					170					175	
Asn	Ser	Ile	Leu	Phe	Ala	Ile	Asn	Ile	Asp	Asn	Lys	Asp	Leu	Asn	Gly
			180					185					190		
Gln	Ser	Lys	Phe	Ala	Pro	Thr	Val	Ser	Asp	Leu	Leu	Lys	Glu	Ser	Thr
		195					200					205			
Gln	Asn	Val	Thr	Ser	Leu	Leu	Lys	Glu	Ser	Thr	Gln	Gly	Val	Ser	Ser
		210				215					220				
Leu	Phe	Arg	Glu	Ile	Thr	Ala	Ser	Ser	Ala	Val	Ser	Ile	Leu	Ile	Lys
225					230					235					240
Pro	Glu	Gln	Glu	Thr	Asp	Pro	Cys	Leu	Ser	Cys	Pro	Gly	Met	Ser	Val
				245					250					255	
Leu	Met	Pro	Asn	Ala	Lys	Arg	Ser	Gly	Arg	Arg	Lys	Arg	Lys	Xaa	Pro
			260					265					270		
Thr	Xaa	Ser	His	Leu	Met	Met	Arg	Lys	Met	Ser	Arg	Thr	Leu	Gly	Thr
		275					280					285			
Cys	Leu	Lys	Arg	His	Leu	Gly	Gln	Gly	Arg	Ala	Gln	Arg	Thr	Thr	Pro
	290					295					300				
Thr	Ala	Pro	Leu	Ser	Ile	Ser	Cys	Pro	Pro	Leu	Lys	Ala	Pro	Ser	Gly
305					310					315					320
Leu	Thr	Pro	Met	Glu	Ser	Glu	Gln	Gln	Leu	Met	Glu	Asn	Xaa	Phe	Pro
				325					330					335	
Val	Phe	Glu	Arg	Gly	Val	Trp	Val	Pro	Glu	Ala	Xaa	Cys	Glu	Lys	His
			340					345					350		
Arg	Xaa	Xaa	Arg	Cys	Gly	Xaa	Lys	Arg	Arg	Xaa	Arg	Val	Trp	Lys	Leu
		355					360					365			
Ile	Arg	Lys	Glu	Ala	Gln	Gly	Pro	Leu	Gly	Val	Ala	Arg	Glu	Ala	Thr
		370				375					380				
Gly	Arg	Glu	His	Leu	Pro	Leu	Pro	Asp	Ala	Gln	Leu	Gly	Ser	Ala	Glu
385					390					395					400
Gly	Ala	Ala	Gln	Xaa	Leu	Arg	His	Pro	Leu	Pro	Cys	Gln	Trp	Arg	Gly
				405					410					415	
Leu	Leu	Gln	Pro	Ser	Arg	Cys	Pro	Pro	Arg	Lys	Pro	Gly	Glu	Arg	Asp
			420					425					430		
Arg	Thr	Arg	Gly	Pro	Arg	Ser	Pro	Gly	Ser	Trp	Thr	Ser	Val	Gln	Cys
		435					440					445			

115
 120
 125
 130
 135
 140
 145
 150
 155
 160
 165
 170
 175
 180
 185
 190
 195
 200
 205
 210
 215
 220
 225
 230
 235
 240
 245
 250
 255
 260
 265
 270
 275
 280
 285
 290
 295
 300
 305
 310
 315
 320
 325
 330
 335
 340
 345
 350
 355
 360
 365
 370
 375
 380
 385
 390
 395
 400
 405
 410
 415
 420
 425
 430
 435
 440
 445

[illegible]

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2631 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

CCCCTCTCAC	AGCCCAGGCC	ATCCAAGAGG	GGCTGAGGAA	AGAGCCCATC	CACCGCGTGT	60
CTGCAGCGGA	GCTGGGAGGG	AAGGTGAACC	GGGCACTACA	GCAAGTGGGA	GGTCTGAAGA	120
GCCCTTGAG	GGGAGAATAT	AAAGAACCAA	GACATCCACC	GCCAAATCAA	GCCAATTACC	180
ACCAGACCCT	CCATGCCCAG	CCGAGAGAGC	TTTCGCCAAG	GGCCCCAGGG	CCCCGGCCAG	240
CTGAGGAGAC	AACAGGCAGA	GCCCCTAAGC	TCCAGCCTCC	TCTCCCACCA	GAGCCCCCAG	300
AGCCAAACAA	GTCTCCTCCC	TTGACTTTGA	GCAAGGAGGA	GTCTGGGATG	TGGGAACCCT	360
TACCTCTGTC	CTCCCTGGAG	CCAGCCCCTG	CCAGAAACCC	CAGCTCACCA	GAGCGGAAAG	420
CAACCGTCCC	GGAGCAGGAA	CTGCAGCAGC	TGGAAATAGA	ATTATTCTC	AACAGCCTGT	480

5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

CCCAGCCATT	TTCTCTGGAG	GAGCAGGAGC	AAATTCTCTC	GTGCCTCAGC	ATCGACAGCC	540
TCTCCCTGTC	GGATGACAGT	GAGAAGAACC	CATCAAAGGC	CTCTCAAAGC	TCGCGGGACA	600
CCCTGAGCTC	AGGCGTACAC	TCCTGGAGCA	GCCAGGCCGA	GGCTCGAAGC	TCCAGCTGGA	660
ACATGGTGCT	GGCCCGGGGG	CGGCCACCG	ACACCCCAAG	CTATTTCAAT	GGTGTGAAAG	720
TCCAAATACA	GTCTCTTAAT	GGTGAACACC	TGCACATCCG	GGAGTTCCAC	CGGGTCAAAG	780
TGGGAGACAT	CGCCACTGGC	ATCAGCAGCC	AGATCCCAGC	TGCAGCCTTC	AGCTTGGTCA	840
CCAAAGACGG	GCAGCCTGTT	CGCTACGACA	TGGAGGTGCC	AGACTCGGGC	ATCGACCTGC	900
AGTGCACACT	GGCCCCTGAT	GGCAGCTTCG	CCTGGAGCTG	GAGGGTCAAG	CATGGCCAGC	960
TGGAGAACAG	GCCCTAACCC	TGCCCTCCAC	CGCCGGCTCC	ACACTGCCGG	AAAGCAGCCT	1020
TCCTGCTCGG	TGCACGATGC	TGCCCTGAAA	ACACAGGCTC	AGCCGTTCCT	AGGGGATYTG	1080
NCCAGCCCCC	CGGCTCARCA	GNTGGGAACC	AGGGCCTCGN	CAGCNAGCNA	AGGTNGGGGG	1140
CAAGCNAGAA	TGCCTCCCAG	GATTTACAN	CCTGAGCCCN	TGCCCCANCC	CTGCTGAADA	1200
AAACAYTNCC	GCCACGTGAA	GAGACAGAAG	GAGGATGGNC	AGGAGTTNNA	CCTYGGGGAA	1260
ACAAAACAGG	GATCTTTNTT	CTGCCCCTGC	TCCAGTNCGA	GTTGGCCTGN	ACCCGCTTGG	1320
ANTCAGTGAC	CATTTGTTGG	CAGANCAGGG	GAGAGCAGCT	TCCAGCCTGG	GTCAGAAGGG	1380
GTGGGCGAGC	CCTTCGGCCC	CTCACCTNC	CAGGCTGCTG	TGNAGAGTGT	CAAGTGTGTA	1440
AGGGNCCCAA	ANCTCAGGNT	TCAGTGCAGA	ACCAGGTNCA	GCAGGTATGC	CCGCCCNGTA	1500
GGTTAANNNG	GGGCCCCTCN	AAACCCCTTG	CCTNGGCCTN	CACCTNGGCC	AGCTCANCCC	1560
CTTTTGGGTG	TAGGGGAAAA	GAATGCCTGA	CCCTGGGAAG	GCTWCCCTGG	TAGAATACAC	1620
CACACTTTTC	AGGTTGTTGC	AACACAGGTC	CTGAGTTGAC	CTCTGGTTCA	GCCAAGGACC	1680
AAAGAAGGTG	TGTAAGTGAA	GTGGTTCTCA	GTNCCCCAGA	CATGTGCCCC	TTTGCTGCTG	1740
GCTACCACTC	TTCCCCAGAG	CAGCAGGCCC	CGAGCCCCTT	CAGGCCCAGC	ACTGCCCCAG	1800
ACTCGCTGGC	ACTCAGTTCC	CTCATCTGTA	AAGGTGAAGG	GTGATGCAGG	ATATGCCTGA	1860
CAGGAACAGT	CTGTGGATGG	ACATGATCAG	TGCTNAAGGN	AAAGCAGCAG	AGAGAGACGY	1920
TCCGGCGCCC	CAGNCCCCAC	TNATCAGTGT	NCCAGCGTGC	TNGGTTNCCC	CAGNAGCACA	1980
GCTNCAGNCA	TCANCACTGA	CACTNCACCC	TNGCCCTGCC	CCTNGGCCAN	GAGGGTACTG	2040
CCGNACGGCA	CTTTGCACNT	CTGATGNACC	TCAAAGCACT	TTCATGGCTN	GCCCTCTNNG	2100
GCAGGGNCAG	GGNCAGGGNC	AGTGACANCT	GTAGGNAGCA	TANGCAANGC	CAGGAGATGG	2160
GGTGNAAGGG	ANCACAGTCT	TGAGCTGTCC	ANCATGCATG	TGACTNCCTC	AAACCTCTTN	2220
NCCAGNATTT	CTCTAAGAAT	AGCANCCCCC	TTNCCCCATT	GCCCCAGCTT	AGCCTCTTCT	2280

CCCAGGGGAG CTANCTCAGG ACACAGTAG CATTAAATCA GCTGTGNAAT CGTCAGGGGG	2340
TGTCTGCTAG CCTCAACCTC CTGGGGCAGG GGACGCCGAG ACTCCGTGGG AGAAGCTCAT	2400
TCCCACATCT TGCCAAGACA GCCTTTNGTC CAGCTGTCCA CATTGAGTCA GACTGCTCCC	2460
GGGAGAGAG CCCC GGCCCC CAGCACATAA AGAACTGCAG CCTTGGTACT GCAGAGTCTG	2520
GGTTGTAGAG AACTCTTTGT AAGCAATAAA GTTTGGGGTG ATGACAAATG TAAAAAAG	2580
GCCTTCGTGG CCTCGAATCA AGCTTATCGA TACCGTCGAC CTCGAGGGGG G	2631

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1253 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

CATTGGAGTC ACGCGGTGGC GGCGCTCTAG AATAGTGGAT CCCC GGGCTG CANGGAATTC	60
GATTCGAGCC CACGAAGGCC CCTTCTTCTG TGGTCGCGGC ACGTTTACAG CCGCAAGCAC	120
CCAGCGGCAG CTGAAGGAGG CTTTGTAGAG GCTCCTGCCC CAGGTGGAGG CGGCCCCGAA	180
GGCCATCCGC GCGCTCAGG TGGAGCGCTA TGTGCCCGAA CACGAGCGAT GCTGCTGGTG	240
CCTGTGCTGC GGCTGTGAGG TGCGGGAACA CCTGAGCCAT GGAAACCTGA CGGTGCTGTA	300
CGGGGGGCTG CTGGAGCATC TGGCCAGCCC AGAGCACAAG AAAGCAACCA ACAAATTCTG	360
GTGGGAGAAC AAAGCTGAGG TCCAGATGAA AGAGAAGTTT CTGGTCACTC CCCAGGATTA	420
TGCGCGATT C AAGAAATCCA TGGTGAAAGG TTTGGATTCC TATGAAGAAA AGGAGGATAA	480
AGTGATCAAG GAGATGGCAG CTCAGATCCG TGAGGTGGAG CAGAGCCGAC AGGAGGTGGT	540
TCGGTCTGTC TTAGAGCCTC AGGCAGTGCC AGACCCAGAA GAGGGCTCTT CAGCACCTAG	600
AAGCTGGAAA GGGATGAACA GCCAAGTAGC TTCCAGCTTA CAGCAGCCCT CAAATTTGGA	660
CCTGCCACCA GCTCCAGAGC TTGACTGGAT GGAGACAGGA CCATCTCTGA CATTCAATTG	720
CCATCAGGAT ATACCAGGAG TTGGTAACAT CCACTCAGGT GCCACACCTC CCTGGATGAT	780
CCAAGATGAA GAATACATTG CTGGGAACCA AGAAATAGGA CCATCCTATG AAGAATTTCT	840
TAAAGAAAAG GAAAAACAGA AGTTGAAAAA ACTCCCCCA GACCGAGTTG GGGCCAACTT	900
TGATCACAGC TCCAGGACCA GTGCAGGCTG GCTGCCCTCT TTTGGGCCGC GTCTGGAATA	960
ATGGACGCCG CTGGCAGTCC AGACATCAAC TCCAAACTG AAGCTGCAGC AATGAAGAAG	1020

CAGTCACATA CAGAAAAAAG CTAATCATGC TCTCTACCAA CTACCATGAG GCTAAAAGCC 1080
 AAAGTCAACC AAACCCCTAT TATACCTTCC ACCCAAATTC TTTATCATTG TCTTTCTTAG 1140
 GAAACAGACA TACTCATTCA TTTGATTTAA TAAAGTTTTA TTTTTCGGCC TTCGTGGCCT 1200
 CGAATCAAGC TTATCGATAC CGTCGACCTC GAGGGGGGGC CGTACCCACT TTT 1253

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 417 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Ile	Gly	Val	Thr	Arg	Trp	Arg	Arg	Ser	Arg	Ile	Val	Asp	Pro	Arg	Ala	1	5	10	15
Ala	Xaa	Asn	Ser	Ile	Arg	Ala	His	Glu	Gly	Pro	Phe	Phe	Cys	Gly	Arg	20	25	30	
Gly	Thr	Phe	Thr	Ala	Ala	Ser	Thr	Gln	Arg	Gln	Leu	Lys	Glu	Ala	Phe	35	40	45	
Glu	Arg	Leu	Leu	Pro	Gln	Val	Glu	Ala	Ala	Arg	Lys	Ala	Ile	Arg	Ala	50	55	60	
Ala	Gln	Val	Glu	Arg	Tyr	Val	Pro	Glu	His	Glu	Arg	Cys	Cys	Trp	Cys	65	70	75	80
Leu	Cys	Cys	Gly	Cys	Glu	Val	Arg	Glu	His	Leu	Ser	His	Gly	Asn	Leu	85	90	95	
Thr	Val	Leu	Tyr	Gly	Gly	Leu	Leu	Glu	His	Leu	Ala	Ser	Pro	Glu	His	100	105	110	
Lys	Lys	Ala	Thr	Asn	Lys	Phe	Trp	Trp	Glu	Asn	Lys	Ala	Glu	Val	Gln	115	120	125	
Met	Lys	Glu	Lys	Phe	Leu	Val	Thr	Pro	Gln	Asp	Tyr	Ala	Arg	Phe	Lys	130	135	140	
Lys	Ser	Met	Val	Lys	Gly	Leu	Asp	Ser	Tyr	Glu	Glu	Lys	Glu	Asp	Lys	145	150	155	160
Val	Ile	Lys	Glu	Met	Ala	Ala	Gln	Ile	Arg	Glu	Val	Glu	Gln	Ser	Arg	165	170	175	
Gln	Glu	Val	Val	Arg	Ser	Val	Leu	Glu	Pro	Gln	Ala	Val	Pro	Asp	Pro	180	185	190	
Glu	Glu	Gly	Ser	Ser	Ala	Pro	Arg	Ser	Trp	Lys	Gly	Met	Asn	Ser	Gln	195	200	205	

Val	Ala	Ser	Ser	Leu	Gln	Gln	Pro	Ser	Asn	Leu	Asp	Leu	Pro	Pro	Ala	
210						215					220					
Pro	Glu	Leu	Asp	Trp	Met	Glu	Thr	Gly	Pro	Ser	Leu	Thr	Phe	Ile	Gly	
225					230					235					240	
His	Gln	Asp	Ile	Pro	Gly	Val	Gly	Asn	Ile	His	Ser	Gly	Ala	Thr	Pro	
				245					250					255		
Pro	Trp	Met	Ile	Gln	Asp	Glu	Glu	Tyr	Ile	Ala	Gly	Asn	Gln	Glu	Ile	
			260					265					270			
Gly	Pro	Ser	Tyr	Glu	Glu	Phe	Leu	Lys	Glu	Lys	Glu	Lys	Gln	Lys	Leu	
		275					280					285				
Lys	Lys	Leu	Pro	Pro	Asp	Arg	Val	Gly	Ala	Asn	Phe	Asp	His	Ser	Ser	
	290					295					300					
Arg	Thr	Ser	Ala	Gly	Trp	Leu	Pro	Ser	Phe	Gly	Pro	Arg	Leu	Glu	Xaa	
305					310					315					320	
Trp	Thr	Pro	Leu	Ala	Val	Gln	Thr	Ser	Thr	Pro	Lys	Leu	Lys	Leu	Gln	
				325					330					335		
Gln	Xaa	Arg	Ser	Ser	His	Ile	Gln	Lys	Lys	Ala	Asn	His	Ala	Leu	Tyr	
			340					345					350			
Gln	Leu	Pro	Xaa	Gly	Xaa	Lys	Pro	Lys	Ser	Thr	Lys	Pro	Leu	Leu	Tyr	
		355					360					365				
Leu	Pro	Pro	Lys	Phe	Phe	Ile	Ile	Val	Phe	Leu	Arg	Lys	Gln	Thr	Tyr	
	370					375					380					
Ser	Phe	Ile	Xaa	Phe	Asn	Lys	Val	Leu	Phe	Phe	Gly	Leu	Arg	Gly	Leu	
385					390					395					400	
Glu	Ser	Ser	Leu	Ser	Ile	Pro	Ser	Thr	Ser	Arg	Gly	Gly	Arg	Thr	His	
				405					410					415		
Phe																

(2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4596 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

AGCGGGGGGA CTGTGCCGTG TGGAACGTGT AGCTGTTGAA GGTGGACTCT GTTACCATTG	60
AGGATGTTTG GAGGATGAGT ATGTGTGGCA GAGGCACACA TAAACAGGCA GAGACCCTTT	120

CCCCCTGCCT	TTCTCCCCCA	ACCCAAGGCT	GACCTGTGTT	CTCCCAGGTC	TGGGATTCTA	180
AGTGACCTGC	TCTGTGTTTG	GTCTCTCTCA	GGATGAGCAC	AAGCCTGGGA	GATGGCAGTG	240
ATGGAAATGG	CCTGCCCAGG	TGCCCCCTGGC	TCAGCAGTGG	GGCAGCAGAA	GGAACTCCCC	300
AAGCCAAAGG	AGAAGACGCC	GCCACTGGGG	AAGAAACAGA	GCTCCGTCTA	CAAGCTTGAG	360
GCCGTGGAGA	AGAGCCCTGT	GTTCTGCGGA	AAGTGGGAGA	TCCTGAATGA	CGTGATTACC	420
AAGGGCACAG	CCAAGGAAGG	CTCCGAGGCA	GGGCCAGCTG	CCATCTCTAT	CATCGCCCAG	480
GCTGAGTGTG	AGAATAGCCA	AGAGTTCAGC	CCCACCTTTT	CAGAACGCAT	TTTCATCGCT	540
GGGTCCAAAC	AGTACAGCCA	GTCCGAGAGT	CTTGATCAGA	TCCCCAACAA	TGTGGCCCAT	600
GCTACAGAGG	GCAAAATGGC	CCGTGTGTGT	TGGAAGGGAA	AGCGTCGCAG	CAAAGCCCCG	660
AAGAAACGGA	AGAAGAAGAG	CTCAAAGTCC	CTGGCTCATG	CAGGAGTGGC	CTTGGCCAAA	720
CCCCTCCCCA	GGACCCCTGA	GCAGGAGAGC	TGCACCATCC	CAGTGCAGGA	GGATGAGTCT	780
CCACTCGGCG	CCCCATATGT	TAGAAACACC	CCGCAGTTCA	CCAAGCCTCT	GAAGGAACCA	840
GGCCTTGGGC	AACTCTGTTT	TAAGCAGCTT	GGCGAGGGCC	TACGGCCGGC	TCTGCCTCGA	900
TCAGAACTCC	ACAAACTGAT	CAGCCCCCTG	CAATGTCTGA	ACCACGTGTG	GAAACTGCAC	960
CACCCCCAGG	ACGGAGGCCC	CCTGCCCCCTG	CCCACGCACC	CCTTCCCCTA	TAGCAGACTG	1020
CCTCATCCCT	TCCCATTCCA	CCCTCTCCAG	CCCTGGAAAC	CTCACCTCT	GGAGTCCTTC	1080
CTGGGCAAAC	TGGCCTGTGT	AGACAGCCAG	AAACCCTTGC	CTGACCCACA	CCTGAGCAAA	1140
CTGGCCTGTG	TAGACAGTCC	AAAGCCCCCTG	CCTGGCCAC	ACCTGGAGCC	CAGCTGCCTG	1200
TCTCGTGGTG	CCCATGAGAA	GTTTTCTGTG	GAGGAATACC	TAGTGCATGC	TCTGCAAGGC	1260
AGCGTGAGCT	CAAGCCAGGC	CCACAGCCTG	ACCAGCCTGG	CCAAGACCTG	GGCAGCACGG	1320
GGCTCCAGAT	CCCGGGAGCC	CAGCCCCAAA	ACTGAGGACA	ACGAGGGTGT	CCTGCTCACT	1380
GAGAAACTCA	AGCCAGTGGA	TTATGAGTAC	CGAGAAGAAG	TCCACTGGGC	CACGCACCAG	1440
CTCCGCCTGG	GCAGAGGCTC	CTTCGGAGAG	GTGCACAGGA	TGGAGGACAA	GCAGACTGGC	1500
TTCCAGTGCG	CTGTCAAAAA	GGTGCGCCTG	GAAGTATTTT	GGGCAGAGGA	GCTGATGGCA	1560
TGTGCAGGAT	TGACCTCACC	CAGAATTGTC	CCTTTGTATG	GAGCTGTGAG	AGAAGGGCCT	1620
TGGGTCAACA	TCTTCATGGA	GCTGCTGGAA	GGTGGCTCCC	TGGGCCAGCT	GGTCAAGGAG	1680
CAGGGCTGTC	TCCCAGAGGA	CCGGGCCCTG	TACTACCTGG	GCCAGGCCCT	GGAGGGTCTG	1740
GAATACCTCC	ACTCACGAAG	GATTCTGCAT	GGGGACGTCA	AAGCTGACAA	CGTGCTCCTG	1800
TCCAGCGATG	GGAGCCACGC	AGCCCTCTGT	GACTTTGGCC	ATGCTGTGTG	TCTTCAACCT	1860
GATGGCCTGG	GAAAGTCCTT	GCTCACAGGG	GACTACATCC	CTGGCACAGA	GACCCACATG	1920

GCTCCGGAGG	TGGTGCTGGG	CACGAGCTGC	GACGCCAAGG	TGGATGTCTG	GAGCAGCTGC	1980
TGTATGATGC	TGCACATGCT	CAACGGCTGC	CACCCCTGGA	CTCAGTTCTT	CCGAGGGGCC	2040
CTCTGCCTCA	AGATTGCCAG	CGAGCCTCCG	CCTGTGAGGG	AGATCCCACC	CTCCTGCGCC	2100
CCTCTCACAG	CCCAGGCCAT	CCAAGAGGGG	CTGAGGAAAG	AGCCCATCCA	CCGCGTGTCT	2160
GCAGCGGAGC	TGGGAGGGAA	GGTGAACCGG	GCACTACAGC	AAGTGGGAGG	TCTGAAGAGC	2220
CCTTGGAGGG	GAGAATATAA	AGAACCAAGA	CATCCACCGC	CAAATCAAGC	CAATTACCAC	2280
CAGACCCTCC	ATGCCCAGCC	GAGAGAGCTT	TCGCCAAGGG	CCCCAGGGCC	CCGGCCAGCT	2340
GAGGAGACAA	CAGGCAGAGC	CCCTAAGCTC	CAGCCTCCTC	TCCCACCAGA	GCCCCCAGAG	2400
CCAAACAAGT	CTCCTCCCTT	GACTTTGAGC	AAGGAGGAGT	CTGGGATGTG	GGAACCCTTA	2460
CCTCTGTCTT	CCCTGGAGCC	AGCCCCTGCC	AGAAACCCCA	GCTCACCAGA	GCGGAAAGCA	2520
ACCGTCCCCG	AGCAGGAACT	GCAGCAGCTG	GAAATAGAAT	TATTCTCAA	CAGCCTGTCC	2580
CAGCCATTTT	CTCTGGAGGA	GCAGGAGCAA	ATTCTCTCGT	GCCTCAGCAT	CGACAGCCTC	2640
TCCCTGTCTG	ATGACAGTGA	GAAGAACCCA	TCAAAGGCCT	CTCAAAGCTC	GCGGGACACC	2700
CTGAGCTCAG	GCGTACACTC	CTGGAGCAGC	CAGGCCGAGG	CTCGAAGCTC	CAGCTGGAAC	2760
ATGGTGCTGG	CCCGGGGGCG	GCCCACCGAC	ACCCCAAGCT	ATTTCAATGG	TGTGAAAGTC	2820
CAAATACAGT	CTCTTAATGG	TGAACACCTG	CACATCCGGG	AGTTCCACCG	GGTCAAAGTG	2880
GGAGACATCG	CCACTGGCAT	CAGCAGCCAG	ATCCCAGCTG	CAGCCTTCAG	CTTGGTCACC	2940
AAAGACGGGC	AGCCTGTTCG	CTACGACATG	GAGGTGCCAG	ACTCGGGCAT	CGACCTGCAG	3000
TGCACACTGG	CCCCTGATGG	CAGCTTCGCC	TGGAGCTGGA	GGGTCAAGCA	TGGCCAGCTG	3060
GAGAACAGGC	CCTAACCCTG	CCCTCCACCG	CCGGCTCCAC	ACTGCCGGAA	AGCAGCCTTC	3120
CTGCTCGGTG	CACGATGCTG	CCCTGAAAAC	ACAGGCTCAG	CCGTTCCCAG	GGGATTGCCA	3180
GCCCCCGGC	TCACAGTGGG	AACCAGGGCC	TCGCAGCAGC	AAGGTGGGGG	CAAGCAGAAT	3240
GCCTCCCAGG	ATTTACACAC	TGAGCCCTGC	CCCACCCTGC	TGAAAAAACA	TCCGCCACGT	3300
GAAGAGACAG	AAGGAGGATG	GCAGGAGTTA	CCTGGGGAAA	CAAAACAGGG	ATCTTTTTCT	3360
GCCCCTGCTC	CAGTCGAGTT	GGCCTGACCC	GCTTGATCA	GTGACCATTT	GTTGGCAGAC	3420
AGGGGAGAGC	AGCTTCCAGC	CTGGGTGAGA	AGGGGTGGGC	GAGCCCTTCG	GCCCCCACC	3480
CTCCAGGCTG	CTGTGAGAGT	GTCAAGTGTG	TAAGGGCCCA	AACTCAGGTT	CAGTGCAGAA	3540
CCAGGTCAGC	AGGTATGCCC	GCCCCTAGGT	TAAGGGGGCC	CTCTAAACCC	CTTGCCCTGGC	3600
CTCACCTGGC	CAGCTCACCC	CTTTTGGGTG	TAGGGGAAAA	GAATGCCTGA	CCCTGGGAAG	3660
GCTCCCTGGT	AGAATACACC	ACACTTTTCA	GGTTGTTGCA	ACACAGGTCC	TGAGTTGACC	3720
TCTGGTTCAG	CCAAGGACCA	AAGAAGGTGT	GTAAGTGAAG	TGGTTCTCAG	TCCCCAGACA	3780

TGTGCCCCTT TGCTGCTGGC TACCACTCTT CCCAGAGCA GCAGGCCCG AGCCCCTTCA 3840
 GGCCCAGCAC TGCCCCAGAC TCGCTGGCAC TCAGTTCCCT CATCTGTAAA GGTGAAGGGT 3900
 GATGCAGGAT ATGCCTGACA GGAACAGTCT GTGGATGGAC ATGATCAGTG CTAAGGAAAG 3960
 CAGCAGAGAG AGACGTCCGG CGCCCCAGCC CCACTATCAG TGTCCAGCGT GCTGGTTCCC 4020
 CAGAGCACAG CTCAGCATCA CACTGACACT CACCCTGCCC TGCCCCTGGC CAGAGGGTAC 4080
 TGCCGACGGC ACTTTGCACT CTGATGACCT CAAAGCACTT TCATGGCTGC CCTCTGGCAG 4140
 GGCAGGGCAG GGCAGTGACA CTGTAGGAGC ATAGCAAGCC AGGAGATGGG GTGAAGGGAC 4200
 ACAGTCTTGA GCTGTCCACA TGCATGTGAC TCCTCAAACC TCTTCCAGAT TTCTCTAAGA 4260
 ATAGCACCCC CTTCCCCATT GCCCCAGCTT AGCCTCTTCT CCCAGGGGAG CTA CTCAGGA 4320
 CTCACGTAGC ATTAAATCAG CTGTGAATCG TCAGGGGGTG TCTGCTAGCC TCAACCTCCT 4380
 GGGGCAGGGG ACGCCGAGAC TCCGTGGGAG AAGCTCATTC CCACATCTTG CCAAGACAGC 4440
 CTTTGTCCAG CTGTCCACAT TGAGTCAGAC TGCTCCCGGG GAGAGAGCCC CGGCCCCCAG 4500
 CACATAAAGA ACTGCAGCCT TGGTACTGCA GAGTCTGGGT TG TAGAGAAC TCTTTGTAAG 4560
 CAATAAAGTT TGGGGTGATG ACAAATGTTA AAAAAA 4596

(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 947 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Met	Ala	Val	Met	Glu	Met	Ala	Cys	Pro	Gly	Ala	Pro	Gly	Ser	Ala	Val
1				5					10					15	
Gly	Gln	Gln	Lys	Glu	Leu	Pro	Lys	Pro	Lys	Glu	Lys	Thr	Pro	Pro	Leu
			20					25					30		
Gly	Lys	Lys	Gln	Ser	Ser	Val	Tyr	Lys	Leu	Glu	Ala	Val	Glu	Lys	Ser
		35					40					45			
Pro	Val	Phe	Cys	Gly	Lys	Trp	Glu	Ile	Leu	Asn	Asp	Val	Ile	Thr	Lys
	50					55					60				
Gly	Thr	Ala	Lys	Glu	Gly	Ser	Glu	Ala	Gly	Pro	Ala	Ala	Ile	Ser	Ile
65					70				75					80	
Ile	Ala	Gln	Ala	Glu	Cys	Glu	Asn	Ser	Gln	Glu	Phe	Ser	Pro	Thr	Phe
			85					90						95	

Ser Glu Arg Ile Phe Ile Ala Gly Ser Lys Gln Tyr Ser Gln Ser Glu
 100 105 110
 Ser Leu Asp Gln Ile Pro Asn Asn Val Ala His Ala Thr Glu Gly Lys
 115 120 125
 Met Ala Arg Val Cys Trp Lys Gly Lys Arg Arg Ser Lys Ala Arg Lys
 130 135 140
 Lys Arg Lys Lys Lys Ser Ser Lys Ser Leu Ala His Ala Gly Val Ala
 145 150 155 160
 Leu Ala Lys Pro Leu Pro Arg Thr Pro Glu Gln Glu Ser Cys Thr Ile
 165 170 175
 Pro Val Gln Glu Asp Glu Ser Pro Leu Gly Ala Pro Tyr Val Arg Asn
 180 185 190
 Thr Pro Gln Phe Thr Lys Pro Leu Lys Glu Pro Gly Leu Gly Gln Leu
 195 200 205
 Cys Phe Lys Gln Leu Gly Glu Gly Leu Arg Pro Ala Leu Pro Arg Ser
 210 215 220
 Glu Leu His Lys Leu Ile Ser Pro Leu Gln Cys Leu Asn His Val Trp
 225 230 235 240
 Lys Leu His His Pro Gln Asp Gly Gly Pro Leu Pro Leu Pro Thr His
 245 250 255
 Pro Phe Pro Tyr Ser Arg Leu Pro His Pro Phe Pro Phe His Pro Leu
 260 265 270
 Gln Pro Trp Lys Pro His Pro Leu Glu Ser Phe Leu Gly Lys Leu Ala
 275 280 285
 Cys Val Asp Ser Gln Lys Pro Leu Pro Asp Pro His Leu Ser Lys Leu
 290 295 300
 Ala Cys Val Asp Ser Pro Lys Pro Leu Pro Gly Pro His Leu Glu Pro
 305 310 315 320
 Ser Cys Leu Ser Arg Gly Ala His Glu Lys Phe Ser Val Glu Glu Tyr
 325 330 335
 Leu Val His Ala Leu Gln Gly Ser Val Ser Ser Ser Gln Ala His Ser
 340 345 350
 Leu Thr Ser Leu Ala Lys Thr Trp Ala Ala Arg Gly Ser Arg Ser Arg
 355 360 365
 Glu Pro Ser Pro Lys Thr Glu Asp Asn Glu Gly Val Leu Leu Thr Glu
 370 375 380
 Lys Leu Lys Pro Val Asp Tyr Glu Tyr Arg Glu Glu Val His Trp Ala
 385 390 395 400
 Thr His Gln Leu Arg Leu Gly Arg Gly Ser Phe Gly Glu Val His Arg
 405 410 415

50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

Met Glu Asp Lys Gln Thr Gly Phe Gln Cys Ala Val Lys Lys Val Arg
 420 425 430
 Leu Glu Val Phe Arg Ala Glu Glu Leu Met Ala Cys Ala Gly Leu Thr
 435 440 445
 Ser Pro Arg Ile Val Pro Leu Tyr Gly Ala Val Arg Glu Gly Pro Trp
 450 455 460
 Val Asn Ile Phe Met Glu Leu Leu Glu Gly Gly Ser Leu Gly Gln Leu
 465 470 475 480
 Val Lys Glu Gln Gly Cys Leu Pro Glu Asp Arg Ala Leu Tyr Tyr Leu
 485 490 495
 Gly Gln Ala Leu Glu Gly Leu Glu Tyr Leu His Ser Arg Arg Ile Leu
 500 505 510
 His Gly Asp Val Lys Ala Asp Asn Val Leu Leu Ser Ser Asp Gly Ser
 515 520 525
 His Ala Ala Leu Cys Asp Phe Gly His Ala Val Cys Leu Gln Pro Asp
 530 535 540
 Gly Leu Gly Lys Ser Leu Leu Thr Gly Asp Tyr Ile Pro Gly Thr Glu
 545 550 555 560
 Thr His Met Ala Pro Glu Val Val Leu Gly Arg Ser Cys Asp Ala Lys
 565 570 575
 Val Asp Val Trp Ser Ser Cys Cys Met Met Leu His Met Leu Asn Gly
 580 585 590
 Cys His Pro Trp Thr Gln Phe Phe Arg Gly Pro Leu Cys Leu Lys Ile
 595 600 605
 Ala Ser Glu Pro Pro Pro Val Arg Glu Ile Pro Pro Ser Cys Ala Pro
 610 615 620
 Leu Thr Ala Gln Ala Ile Gln Glu Gly Leu Arg Lys Glu Pro Ile His
 625 630 635 640
 Arg Val Ser Ala Ala Glu Leu Gly Gly Lys Val Asn Arg Ala Leu Gln
 645 650 655
 Gln Val Gly Gly Leu Lys Ser Pro Trp Arg Gly Glu Tyr Lys Glu Pro
 660 665 670
 Arg His Pro Pro Pro Asn Gln Ala Asn Tyr His Gln Thr Leu His Ala
 675 680 685
 Gln Pro Arg Glu Leu Ser Pro Arg Ala Pro Gly Pro Arg Pro Ala Glu
 690 695 700
 Glu Thr Thr Gly Arg Ala Pro Lys Leu Gln Pro Pro Leu Pro Pro Glu
 705 710 715 720
 Pro Pro Glu Pro Asn Lys Ser Pro Pro Leu Thr Leu Ser Lys Glu Glu
 725 730 735
 Ser Gly Met Trp Glu Pro Leu Pro Leu Ser Ser Leu Glu Pro Ala Pro

6
 5
 4
 3
 2
 1
 0
 1
 2
 3
 4
 5
 6

740

745

750

Ala Arg Asn Pro Ser Ser Pro Glu Arg Lys Ala Thr Val Pro Glu Gln
755 760 765

Glu Leu Gln Gln Leu Glu Ile Glu Leu Phe Leu Asn Ser Leu Ser Gln
770 775 780

Pro Phe Ser Leu Glu Glu Gln Glu Gln Ile Leu Ser Cys Leu Ser Ile
785 790 795 800

Asp Ser Leu Ser Leu Ser Asp Asp Ser Glu Lys Asn Pro Ser Lys Ala
805 810 815

Ser Gln Ser Ser Arg Asp Thr Leu Ser Ser Gly Val His Ser Trp Ser
820 825 830

Ser Gln Ala Glu Ala Arg Ser Ser Ser Trp Asn Met Val Leu Ala Arg
835 840 845

Gly Arg Pro Thr Asp Thr Pro Ser Tyr Phe Asn Gly Val Lys Val Gln
850 855 860

Ile Gln Ser Leu Asn Gly Glu His Leu His Ile Arg Glu Phe His Arg
865 870 875 880

Val Lys Val Gly Asp Ile Ala Thr Gly Ile Ser Ser Gln Ile Pro Ala
885 890 895

Ala Ala Phe Ser Leu Val Thr Lys Asp Gly Gln Pro Val Arg Tyr Asp
900 905 910

Met Glu Val Pro Asp Ser Gly Ile Asp Leu Gln Cys Thr Leu Ala Pro
915 920 925

Asp Gly Ser Phe Ala Trp Ser Trp Arg Val Lys His Gly Gln Leu Glu
930 935 940

Asn Arg Pro
945

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide PCR primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

CAGGATCCTC ATGGCTGCAG CTAGCGTGAC

30

(2) INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide PCR primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GGTCGACTTA GAGCCCTGTC AGGTCCACAA TG

32

(2) INFORMATION FOR SEQ ID NO: 10:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide probe"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

GATGCCATTG GGGATTTCCT CTTT

24

(2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide probe"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CAGTAAAGAG GAAATCCCCA ATGG

24